Chapter 14

REGIONAL FOREST LAND COVER CHARACTERISATION USING MEDIUM SPATIAL RESOLUTION SATELLITE DATA

Chengquan Huang, Collin Homer, and Limin Yang
USGS EROS Data Center, Raytheon, Sioux Falls, South Dakota, 57198, USA

1. INTRODUCTION

Increasing demands on forest resources require comprehensive, consistent and up-to-date information on those resources at spatial scales appropriate for management decision-making and for scientific analysis. While such information can be derived using coarse spatial resolution satellite data (e.g. Tucker et al. 1984; Zhu and Evans 1994; Cihlar et al. 1996; Cihlar et al., Chapter 12), many regional applications require more spatial and thematic details than can be derived by using coarse resolution imagery. High spatial resolution satellite data such as IKONOS and Quick Bird images (Aplin et al. 1997), though usable for deriving detailed forest information (Culvenor, Chapter 9), are currently not feasible for wall-to-wall regional applications because of extremely high data cost, huge data volume, and lack of contiguous coverage over large areas. Forest studies over large areas have often been accomplished using data acquired by intermediate spatial resolution sensor systems, including the Multi-Spectral Scanner (MSS), Thematic Mapper (TM) and the Enhanced Thematic Mapper Plus (ETM+) of Landsat, the High Resolution Visible (HRV) of the Systeme Pour l'Observation de la Terre (SPOT), and the Linear Image Self-Scanner (LISS) of the Indian Remote Sensing satellite. These sensor systems are more appropriate for regional applications because they can routinely produce spatially contiguous data over large areas at relatively low cost, and can be used to derive a host of forest attributes (e.g. Cohen et al. 1995; Kimes et al.

1999; Cohen et al. 2001; Huang et al. 2001; Sugumaran 2001). Of the above intermediate spatial resolution satellites, Landsat is perhaps the most widely used in various types of land remote sensing applications, in part because it has provided more extensive spatial and temporal coverage of the globe than any other intermediate resolution satellite. Spatially contiguous Landsat data have been developed for many regions of the globe (e.g. Lunetta and Sturdevant 1993; Fuller et al. 1994b; Skole et al. 1997), and a circa 1990 Landsat image data set covering the entire land area of the globe has also been developed recently (Jones and Smith 2001). An acquisition strategy aimed at acquiring at least one cloud free image per year for the entire land area of the globe has been initiated for Landsat-7 (Arvidson et al. 2001). This will probably ensure the continued dominance of Landsat in the near future.

Extracting forest information from Landsat imagery has been a vigorous research activity over the past 30 years of Landsat history. Early forest applications used both digital and analogue methods to analyse satellite imagery. Images were often digitally enhanced for printing, with actual interpretation done on the hard-copy print. With the evolving need for more efficient and comprehensive analysis, a wide range of digital methods were developed (Townshend 1992; Hall et al. 1995), many of which were tested within local areas covered by single Landsat scenes. Rapid development in computer hardware and software over the last decade provided the computing capacity for deriving forest information from multiple satellite scenes. (Here, and throughout this Chapter, a scene refers to an area covered by a Landsat path/row, while an image refers to a specific image acquisition.) For example, Bauer (1994) mapped seven forest classes in five north-eastern Minnesota counties using six TM scenes. At national scales, over 20 land cover classes, including three forest classes, were mapped using TM images for Great Britain (Fuller et al. 1994a) and the conterminous U.S. (Vogelmann et al. 2001).

Use of a single TM image for forest studies can be a very complex process. Some of the challenges include radiometric and geometric correction, impact of topography and the atmosphere on image quality, and diversity and spatial heterogeneity of land cover, among others (Jensen 1986). Deriving forest information from multiple scenes further compounds these issues. For example, cloud cover often makes it difficult to obtain usable images for an entire study area within a specific time window. Another challenge is among-scene variability arising from differences in atmospheric condition, viewing and illumination geometry, vegetation phenology and soil moisture content. As a result, information extraction methods that work well in a single-scene application may fall apart in multi-scene applications, or may be less effective when trained on one scene and applied to neighbouring scenes (Pax-Lenney et al. 2001). A third challenge is lack of reliable reference data sets. Such data sets are required for training many classification algorithms (see Franklin et al., Chapter 10) and for accuracy assessment (Czaplewski, Chapter 5). Because reference data are expensive to collect, at regional scales they are often compiled from different sources, provided such sources exist. Regardless, reference data sets from different sources are often collected by different parties at different times using different methods, and often have varying levels of reliability (DeFries and Townshend 1993). Use of such reference data sets may result in inconsistencies and varying reliability in derived data products.

In this Chapter, we will address these challenges through two case studies – the land cover mapping project of the Utah Gap Analysis Program (GAP) (Homer et al. 1997) and a pilot study of the Multi-Resolution Land Characteristics (MRLC) 2000 program, which for simplicity are referred to as Utah GAP land cover program and MRLC 2000 pilot study, respectively. While the unsupervised clustering approach employed in the Utah GAP land cover program has been used to develop many large area classifications (e.g. Cihlar et al. 1996; Cohen et al. 1998; Vogelmann et al. 1998), the supervised classification tree method used in the MRLC 2000 pilot study is gaining popularity due to its promising performance in regional and global applications (e.g. Friedl et al. 1999; Hansen et al. 2000). In the following sections, we first present the two case studies, with more emphasis on the MRLC 2000 pilot study. Results of the two case studies are then evaluated using an independent reference data set, following which some of the major issues on regional forest land cover characterisation using medium spatial resolution satellite data are discussed.

2. THE UTAH GAP LAND COVER PROGRAM

2.1 Background

In 1990, the U.S. Geological Survey (USGS) Gap Analysis Program (GAP) was established to map terrestrial vertebrate species and evaluate their protection status on the land where they occur throughout the United States (Edwards et al. 1993; Scott et al. 1993). Central to this analysis was a vegetation cover-type map which, when linked to wildlife habitat relation models, predicts the spatial distribution of animal species. Because no regional land cover information existed at the time when the program started, one of the requirements of this program was to develop a state-wide vegetation cover-type map from Landsat data. This case study describes the development of this vegetation map for the state of Utah.

2.2 Image Standardisation and Mosaic Creation

Large study areas requiring the spatial resolution of Landsat TM data invariably cover multiple scenes. Analysis and classification of multiple scenes can be carried out either on individual scenes or multi-scene mosaics. Single-scene classification potentially offers better accuracy because of reduced pixel sample size and spectral variability. However, classifying single scenes independently within a multi-scene region can require a greater investment in time, training data collection, and subsequent edge-matching than the multi-scene mosaic approach. On the other hand, the mosaic approach has a possible disadvantage of increased within-class spectral variability and potentially higher confusion between spectrally similar cover types. For this application, the mosaic approach offered the best solution to our mapping objectives.

Spatially, 14 Landsat TM scenes cover Utah. Twenty-four images were required to provide a complete, cloud free mosaic for a single summer season, including 14 primary base scenes and ten secondary cloud patch scenes. The primary base scenes were acquired between June and August of 1988 and 1989. The dates for the additional ten images used for cloud patching varied from 1984 to 1993, but were all in the summer growing season.

A two-step approach of atmospheric standardisation and histogram adjustment was chosen to normalise image-to-image variations. First, the image acquired at path 37, row 33 was chosen as a "master" image because of its central location in the state (allowing maximum overlay with adjacent scenes) and because it covered a significant range of the ecological conditions likely to be encountered in the state. This image was adjusted for atmospheric haze by plotting each of the reflective spectral bands against the middle infrared band 7 (2.08 to 2.35 μm) as described in Jensen (1986). Then, a histogram adjustment method based on histogram bias (i.e., histogram shape is maintained but relative position is altered) was used to normalise among-scene variations due to the additive components of atmospheric effects. This method does not alter within-slave-scene unique radiometric characteristics, allowing for recognition of localised phenomena in digital classification. Selected areas of overlap between master and slave were compared band-by-band, and the average difference was calculated for each band. Band-by-band differences from the overlap sample areas were used as bias values to adjust radiometrically the slave to the master. Once a slave image was radiometrically matched to the master, it became a master for its adjacent scenes. This method effectively reduced the among-scene variations in this study area, resulting in a near seamless mosaic (Colour Plate 16).

To reduce the spectral variability within individual classes and possible confusion between spectrally similar but ecologically different cover types, the image mosaic was segmented into three ecoregions, i.e., Wasatch-Uinta, Colorado Plateau, and Northern Great Basin. The Wasatch-Uinta ecoregion is characterised by high mountains and plateaux containing typical rocky mountain flora such as evergreen and deciduous forest. The Colorado plateau is characterised by lower elevation canyon lands, plateaux and buttes supporting arid and semi-arid shrub, grass, and woodlands. Typical landscapes in the Northern Great Basin ecoregion include mountain ranges and broad valleys trending north to south (Colour Plate 16, after Omernik (1987)).

2.3 Training Data

Though an unsupervised clustering approach was employed in this case study, training data were needed to link spectral clusters to vegetation cover types. The required training data were collected through interpreting low altitude aerial photos and conducting field work, with field location determined using Global Positioning System (GPS) units. A total of 657 field training sites were used in the Wasatch-Uinta ecoregion, of which 356 (53 percent) were collected using GPS units in the field, 221 (35 percent) using low-altitude aerial photographs, and 80 (12 percent) using both methods. Of the 518 training sites in the Colorado Plateau ecoregion, 422 (81 percent) were GPS located, 59 (12 percent) were photo-interpreted, and 37 (7 percent) were collected using both methods. In the Northern Basin and Range ecoregion, 490 training sites (86 percent) were GPS based, 26 (4 percent) were photo-interpreted, and 57 (10 percent) were collected using both methods, totalling 573 training sites.

2.4 Classification and Modelling

The three ecoregions were subset from the state image mosaic and processed using the ISODATA algorithm implemented in the ERDAS image processing package to generate unsupervised spectral clusters (Tou and Gonzalez 1974). Before clustering, agricultural and urban areas were masked from the image using an existing Geographic Information System database to further reduce spectral variability. A total of 125 clusters were initially generated in the Wasatch-Uinta ecoregion, with 150 each in the Colorado Plateau and Northern Basin and Range. A minimum-distance-to-the-means classification algorithm was used to assign individual pixels to a spectral class.

Cover-type modelling consisted of two phases: statistical association of spectral classes with cover-types and post-classification ecological modelling based on ancillary information. The first phase of modelling included two sets of summary statistics generated from the weighted training polygon values: the proportion of each cover-type weighting ordered by spectral class and spectral class polygon weighting values ordered by cover type. These two sets of summary statistics were used in concert during ancillary modelling to provide balance between possible commission and omission errors.

The second phase of modelling incorporated ancillary data, including 3-arc-second resolution digital elevation, slope, aspect, and vegetation cover-type range polygons, to clarify cover-type associations by using post-classification stratification (Fleming and Hoffer 1979; Hutchinson 1982). Polygons delineating the general distribution of vegetation cover-type were developed from existing literature and maps, and were used to limit the geographic extent of some cover types. All localised ancillary parameters detailed in literature and field work were standardised to regional scales before being used in modelling. An intensive effort was made to ensure as much objectivity as possible in generating spectral class/ancillary data models. This second phase modelling was extensive for some spectral classes.

2.5 Results

A 38-class land cover map with sixteen forest classes was developed in this case study (Homer et al. 1997). An accuracy assessment based on a mixed sampling design that considered statistical validity, accessibility and efficacy yielded an overall map accuracy of 75.3 % for the entire state (Edwards et al. 1998). Further evaluation of this classification using an independent reference data set will be presented in a later section.

3. MRLC 2000 PILOT STUDY

3.1 Background

The MRLC consortium was initiated in early 1990s to address the need for consistent national and regional land cover data for the United States (Loveland and Shaw 1996). Through this consortium, a 1992-vintage National Land Cover Dataset (NLCD 1992) was developed for the conterminous United States (Vogelmann et al. 2001). MRLC is currently developing a second generation land cover product, NLCD 2000, using 2000-vintage Landsat-7 ETM+ images and relevant ancillary data. The guiding principles in designing NLCD 2000 included the need to: 1) develop methods that are as objective, consistent and repeatable as possible for generating standardised land cover products, 2) constrain methods to be simple, efficient and transferable to others, 3) develop land cover products flexible enough to meet the potentially diverse requirements of multiple users, 4) provide users with increased access to intermediate database products and derivatives enabling local application, and 5) maintain reasonable compatibility with NLCD 1992. The purpose of this MRLC 2000 pilot study is to develop a prototype procedure that follows the above guiding principles and is efficient and robust for use in all regions of the U.S. The MRLC 2000 classification scheme consists of over 20 land cover classes (Homer et al. 2002). We will examine only the forest classes in this Chapter.

3.2 Data and methods

The overall procedure of this pilot study consists of pre-processing of satellite imagery, ancillary data and reference data, classification using a decision tree method, and accuracy assessment using both cross-validation and independent test data sets (Figure 14-1). The study area primarily covered the Rocky Mountains of Utah, extending from the Cache National Forest, located north of Salt Lake City, to Zion National Park in the south. A large part of this study area overlapped with the Wasatch-Uinta ecoregion of the Utah GAP land cover mapping program (Colour Plate 16).

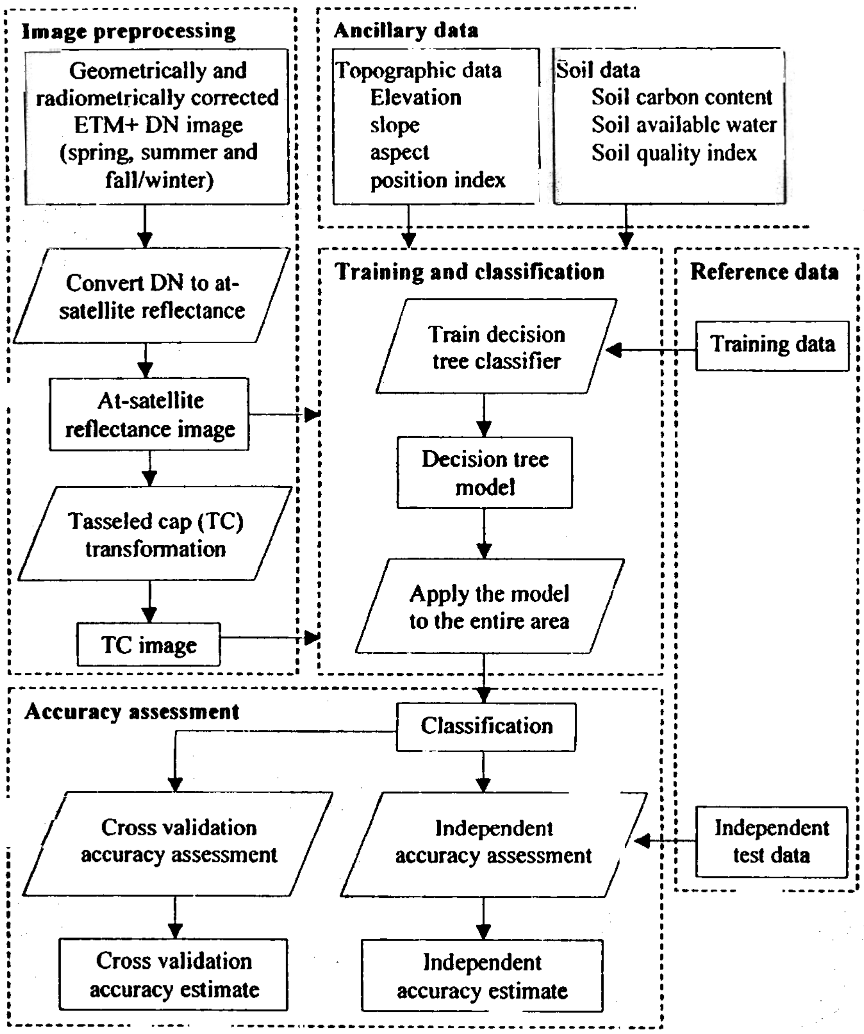

Figure 14-1. A flowchart of data and methods used in the MRLC 2000 pilot study.

3.2.1 Satellite imagery and ancillary data

Nine ETM+ scenes were required to cover the entire study area (Table 14-1). For each scene, near cloud free ETM+ images acquired on three dates between 1999 and 2001 were used to capture vegetation dynamics over a growing season and to maximise land cover type separability. Image selection was based on vegetation greenness profiles defined by a multi-year normalized difference vegetation index data set derived from the Advanced

Very High Resolution Radiometer (Yang et al. 2001). Two additional images acquired in the summer of 2000 were used to patch the clouds seen in two summer leaf-on images. All images were geometrically and radiometrically corrected using standard methods at the USGS EROS Data Center (Irish 2000). Terrain correction using the USGS 1-arc second National Elevation Dataset was performed to improve geolocation accuracy. To reduce among-scene variability due to different illumination geometry (Colour Plate 17A), raw digital numbers were converted to at-satellite reflectance for the 6 reflective bands, and to at-satellite temperature for the thermal band according to Markham and Barker (1986) and the Landsat-7 Science Data User's Handbook (Irish 2000). Colour Plate 17B shows that a large portion of the among-scene variations were removed using this method (also demonstrated in Huang et al. 2002b). All 7 bands of the images were resampled to a 30 m spatial resolution. Tasseled-cap brightness, greenness and wetness were calculated using at-satellite reflectance based coefficients (Huang et al. 2002b). Mosaics of the study area were developed using the at-satellite reflectance images and corresponding tasseled-cap images. Colour Plate 17C shows the summer leaf-on image mosaic.

Table 14-1. Landsat ETM+ images used in the MRLC 2000 pilot study. The unit for sun elevation is degree.

	Spring		Summer		Fall/winter	
Landsat path/row	Acquisition date	Sun elevation	Acquisition date	Sun elevation	Acquisition date	Sun elevation
36/32	04/28/2000	58	06/15/2000	65	10/19/1999	37
37/31	05/08/2001	60	06/06/2000	64	09/10/2000	49
37/32	05/08/2001	60	06/06/2000	65	10/10/1999	40
37/33	05/05/2000	61	06/06/2000	65	10/10/1999	41
37/34	05/05/2000	62	06/06/2000	66	10/10/1999	43
38/31	05/28/2000	63	06/29/2000	64	10/03/2000	41
38/32	05/28/2000	64	08/14/1999	57	10/17/1999	38
38/33	04/26/2000	59	07/31/2000	61	11/02/1999	34
38/34	04/26/2000	60	07/31/2000	61	11/02/1999	35

Ancillary data included the USGS 1-arc second National Elevation Dataset and three derivatives, i.e., slope, aspect and a topographic position index characterising a pixel's position relative to ridges and valleys. In addition, three soil attributes, i.e. available water capacity, soil carbon content and a soil quality index, were derived from the State Soil Geographic (STATSGO) Data Base (USDA 1991). All ancillary data layers were rasterized or resampled to have a spatial resolution of 30 m.

3.2.2 Reference Data Sets

Two reference data sets were available to this study. One was collected through the Forest Inventory and Analysis (FIA) program of the U.S. Forest Service in mid-1990. Through intensive field work, the FIA program provided detailed forest attributes at individual tree, sub-plot and plot levels. Covering an area of about 2 × 2 ETM+ pixels, an FIA plot consists of 4 to 5 sub-plots, each with a radius of about 7.3 meter. Considering the pixel size of ETM+ imagery and possible geolocation errors, only plot level data were deemed appropriate for use with the ETM+ imagery. There were 3037 FIA plot in this study area. Each FIA plot was classified using two classification schemes, forest/non-forest and a 4-class scheme. In addition, the 1852 forest plots were also classified using a forest type group scheme. Both the 4-class scheme and the forest type group scheme will be listed in the following section.

The other reference data set consisted of 1295 field points collected by the Fire Science Lab of the Rocky Mountain Research Station (RMRS) of the U.S. Forest Service in late 1990s. These points were distributed within the mid-southern portion of the study area. The collected information allowed the labelling of each point using the forest/non-forest and 4-class schemes. This data set was used to partially evaluate the classification results developed using FIA plot data.

3.2.3 Classification schemes

As mentioned in the previous section, three classification schemes were considered in this study: forest/non-forest, a 4-class scheme consisting of non-forest, deciduous, evergreen, and mixed, and a forest type group scheme. The FIA program requires a forest/non-forest map to implement a stratified sampling of forested land in order to produce accurate estimates of forest attributes (McRoberts et al. 2002). Deciduous, evergreen and mixed forest are the common forest categories in many regional land cover classification systems, including the MRLC 2000 classification scheme. Information on forest type group is required for species conservation planning, fire management and many other applications. Based on the FIA plot data, there were 9 major forest type groups in this area: pinyon/juniper, douglas-fir, ponderosa pine, fir/spruce/mountain hemlock, lodgepole pine, other western softwoods, aspen/birch, western oak and other western hardwoods.

3.2.4 Decision Tree Classifier

The ample reference data points available to this study made it possible to employ a supervised approach for deriving forest land cover classifications (Richards 1993). The popular supervised classification algorithms include the maximum likelihood classifier, neural networks and decision tree methods (Townshend 1992; Hall et al. 1995; Franklin et al., Chapter 10). The decision tree method was chosen for this study because it 1) is non-parametric and therefore independent of the distribution of class signature, 2) can handle both continuous and non-continuous variables, 3) generates interpretable classification rules, 4) is fast to train and is often as accurate as and sometimes more accurate than, many other classifiers (Hansen et al. 1996; Huang et al. 2002a). Tree classifiers have been used to develop land cover classifications at regional to global scales (e.g. Friedl and Brodley 1997; DeFries et al. 1998; Friedl et al. 1999; Hansen et al. 2000). The decision tree program used in this study, C5, employs an information gain ratio criterion in tree development and pruning. A general description of the functions of this program is given in a tutorial available at http://www.rulequest.com/see5-unix.html. A detailed description of an earlier version of this program, C4.5, was provided by Quinlan (1993).

One of the useful functions of C5 is boosting, a technique designed to improve classification accuracy (Bauer and Kohavi 1998). With this function, the program develops a sequence of decision trees, with each subsequent one trying to fix the misclassification errors in the previous tree. Each decision tree makes an independent prediction, and the final prediction is a weighted vote of the predictions of all trees. This function often improves classification accuracy by 5 % to 10 % (e.g. Friedl et al. 1999; Chan et al. 2001).

3.2.5 Accuracy assessment

Accuracy estimates of the classifications were derived in two ways. One was to use the independent test data set collected by the Fire Science Lab of the U. S. Forest Service RMRS. The other was to use a cross-validation function of C5. Cross-validation is designed to derive prompt accuracy estimates even when only a limited number of reference data samples are available for both training and accuracy assessment (Henery 1994). For an N-fold cross-validation, the training data set is divided into N equal-sized subsets. Accuracy estimates are derived by using each subset to evaluate the classification developed using the remaining training samples. The mean accuracy and its standard error represent those of the classification developed using all reference samples.

3.3 Results

Two classifications, one with and the other without the use of boosting, were developed for each of the three classification schemes – forest/non-forest, 4-class and forest type group, using FIA plot data and Landsat-7 ETM+ images. Classification accuracies derived using cross-validation and the independent reference data set collected by the Fire Science Lab of RMRS are reported in Table 14-2. With the boosting function of the C5 program, overall accuracies of around 80 % were achieved for the forest/non-forest and the 4-class classifications and about 65 % for forest type group classification. The boosting function improved classification accuracy by about 2 to 9 percent in absolute values. Similar improvements using the boosting function have been reported in other studies (e.g. Chan et al. 2001). These classifications were visually evaluated by field crew members of RMRS and the Utah GAP Analysis program of Utah State University. Both parties agreed that these classifications were reasonably accurate.

Table 14-2. Classification accuracy estimates for the MRLC 2000 pilot study. The units for both accuracy and standard error are percent (%).

Classification level	Forest/non-forest		4-class		Forest type group	
	Accuracy	Std. Error	Accuracy	Std. Error	Accuracy	Std. Error
Cross-validation						
- Without boosting	80.4	0.4	78.0	0.4	56.6	0.9
- With boosting	82.7	0.4	81.2	0.6	65.8	1.2
Independent assessment						
- Without boosting	75.7	-	75.3	-	-	-
- With boosting	79.0	-	83.4	-	-	-

4. COMPARISON OF THE TWO CASE STUDIES

Using an independent reference data set provided by the FIA program of U.S. Forest Service RMRS, we were able to compare the classifications derived through the Utah GAP land cover program and the MRLC 2000 pilot study in the Wasatch-Uinta area, where the two study areas overlapped. This reference data set consisted of 68,358 points regularly spaced at a 1km interval in both the east-west and south-north directions. Based on aerial photos acquired in the 1980s, each point was labelled with one of the following classes: non-forest, conifer, pinyon/juniper, aspen and other hardwoods. This reference data set allowed an independent assessment of the classifications developed in the two case studies by aggregating those classifications to this 5-class level as well as to the 4-class and forest/non-forest levels. The derived overall accuracies as well as both user's and producer's accuracies are reported in Table 14-3.

Table 14-3. Accuracies (%) estimated using 68,358 photo-interpreted points for classifications developed in the two case studies. Accuracies for the mixed class were unavailable because the reference data did not have a mixed class. MRLC 2000 classifications were developed using C5's boosting function.

	User's		Producer's		Overall	
	Utah GAP	MRLC 2000	Utah GAP	MRLC 2000	Utah GAP	MRLC 2000
forest/non-forest					70.7	78.4
non-forest	62.3	69.7	68.7	75.9		
forest	77.3	86.2	72.0	82.1		
4-class					62.8	71.3
non-forest	62.3	69.7	68.7	82.1		
deciduous	55.6	62.4	49.1	52.3		
evergreen	68.3	77.0	63.2	69.2		
mixed	-	-				
5-class					58.5	66.5
non-forest	62.3	69.7	68.7	82.1		
conifer	63.5	71.7	56.4	64.9		
other hardwoods	42.4	53.6	35.9	29.5		
aspen	44.0	48.7	46.2	56.7		
pinyon/juniper	58.2	66.3	56.2	59.2		

The overall accuracies of the classifications at the forest/non-forest, 4-class and 5-class levels were 70.7 %, 62.8 % and 58.5 % for the Utah GAP land cover program, and 78.4 %, 71.3 % and 66.5 % for the MRLC 2000 pilot study, respectively. The differences in overall accuracy between the two sets of classifications were about 8 % at all three levels. Classifications of the MRLC 2000 pilot study also had class specific user's accuracies of about 4 % - 11 %, and producer's accuracies of about 3 % - 13 % (except for the other hardwoods class) higher than the Utah GAP classifications. With the exception of the 4-class classification, the overall accuracies of the other two classifications of the MRLC 2000 pilot study were comparable to those estimated through cross-validation (Table 14-2, with boosting) or using the independent test data set collected by the Fire Science Lab, suggesting that these estimates were not significantly biased from each other. Colour Plate 18 shows a window of the two classifications at the 5-class level and the summer leaf-on ETM+ image used in the MRLC 2000 pilot study. The overall accuracy of the 4-class map was about 10 % lower than the cross-validation estimate (Table 14-2, with boosting) and the one derived using the independent test data set collected by the Fire Science Lab. This might be partially due to the lack of a mixed class in the photo-interpreted reference data set.

The higher classification accuracies of the MRLC 2000 pilot study compared to the Utah GAP land cover program can be attributed to at least two factors. One is use of multi-temporal imagery, which often yields better classification accuracies than using single-date imagery (e.g. Coppin and Bauer 1994; Lunetta and Balogh 1999). Whereas only a single-date summer leaf-on image mosaic was used in the Utah GAP land cover program, image mosaics of three dates representing spring, summer and fall/winter were used in the MRLC 2000 pilot study. The other factor is use of a high quality training data set – FIA plot data. This data set consists of points collected following a probability-based sampling design. Each point was labelled according to intensive field work, and was revisited periodically. Collected nation-wide, this data set is highly valuable for use with intermediate spatial resolution satellite imagery in regional forest studies. To ensure the confidentiality, security and integrity of the data points, however, use of this data set should be arranged under security agreements.

5. DISCUSSION

The two case studies presented in this work demonstrate the feasibility of extracting forest information at regional scales using multiple Landsat scenes as a single mosaic. This information extraction process consists of a sequence of steps, including normalising among-scene variations independent of land surface conditions, selecting appropriate classification algorithms and validating classification results. As a result of the increased acquisition capacity of many local Landsat receiving stations, together with the high priority being given to the acquisition of global coverage by Landsat-7 (Arvidson et al. 2001), data availability has become less problematical in regional Landsat applications. The methods used in each step, however, can affect the reliability, efficiency and consistency of deriving forest information from Landsat data.

5.1 Normalisation of among-scene variability

Images of neighbouring scenes, especially of neighbouring Landsat paths, are often acquired on different dates. They can differ in atmospheric conditions, illumination geometry, and vegetation phenology, resulting in increased within-class signature variability and reduced among-class separability when these images are analysed as a mosaic. It is therefore desirable to normalise such among-scene variations before information extraction. Designed to retrieve surface reflectance from digital number, physically based atmospheric correction algorithms are, in theory, preferable for standardising the impact of the atmosphere and illumination geometry. However, use of available atmospheric correction algorithms on Landsat imagery over large areas has very limited success, partially because many required parameters concerning *in situ* atmospheric conditions are often not available or cannot be derived reliably (Cohen et al. 2001). Use of pseudo-invariant objects whose reflective properties remain relatively stable may provide a partial solution (Schott et al. 1988), provided enough pseudo-invariant objects can be identified in the overlap areas of neighbouring scenes. The histogram bias adjustment method introduced in the Utah GAP case study is similar to a dark object subtraction approach to atmospheric correction. One of the limitations of this approach is that it only normalises additive components of atmospheric effects (Teillet and Fedosejevs 1995). It can not handle non-additive components properly. The at-satellite reflectance method described in the MRLC 2000 pilot study effectively normalized the impact of illumination geometry, a non-additive component. For clear and near-cloud-free images, use of this method alone may generate satisfactory image mosaics. However, when varying hazy conditions exist among the images, applying the at-satellite reflectance method followed by the histogram bias adjustment method may substantially improve the quality of an image mosaic.

No efforts were made to normalise the among-scene variations arising from differences in vegetation phenology in the two case studies. Because such variations are functions of a number of factors, including vegetation type, agricultural activity, and perhaps topographically-induced soil moisture availability, there are currently no practical solutions to this problem. In the MRLC 2000 pilot study, we tried to address this problem by using scene identification number as an input to the decision tree program, and effectively removed some seamlines seen in the image mosaic from the derived classifications (Colour Plate 18).

5.2 Classification algorithm selection

The two case studies represent two different approaches to land cover classification, one supervised and the other unsupervised. While classifications developed in the MRLC 2000 pilot study had higher overall accuracies than those developed in the Utah GAP land cover program, it was not clear if the differences were partially due to selection of classification algorithms. As discussed in a previous section, two other factors, i.e., use of multi-temporal images and better reference data, may also have contributed to this. Whether to use a supervised or unsupervised method in a specific project depends on many factors. Supervised methods generally require substantial amounts of reliable reference data in order to be trained adequately. Because of this, Richards (1993) suggested that the supervised maximum likelihood classifier might be more time demanding than an unsupervised approach. However, results from the two case studies suggest that, with reliable and up-to-date reference data sets like the FIA plot data readily available, a supervised method is often more efficient and cost-effective than an unsupervised method. We estimated that, for land cover mapping with similar thematic content, the MRLC 2000 pilot study took less than one third of the effort the Utah GAP program took. Supervised methods will also probably be more efficient for remapping efforts, because for most areas the majority of the training points used in a previous mapping effort are not likely to change from one cover type to another during the mapping interval. The unchanged training points should be reusable after being identified through quality check. In addition to the advantages of the decision tree program discussed in the MRLC 2000 pilot study, this program can also produce instant accuracy estimates for the classification being developed, which may be relatively unbiased if the training data points are collected following a probability-based sampling design and are not spatially auto-correlated.

Unfortunately, for many large area forest land cover mapping activities, there are often insufficient reference data samples, and developing a large reference data set with adequate samples may not always be feasible for many practical reasons, including insufficient resources, time constraints and access problems. While it is often difficult to use supervised approaches in such cases, the Utah GAP land cover program and other studies (e.g. Vogelmann et al. 1998) have demonstrated that an unsupervised approach can yield satisfactory results.

Accuracy assessment

Validation of classification results at a regional scale is a considerable task. A statistically valid accuracy assessment can take tremendous amounts of resources and time. Even if the required resources are available, it often take several months to several years to produce valid accuracy values after a classification is developed. To avoid this problem, it is highly recommended that the accuracy assessment be planned during the design phase of a mapping project. An accuracy assessment plan should include three components: a probability-based sampling design, a protocol for labelling reference data points, and an appropriate statistical procedure for deriving accuracy estimates (Czaplewski, Chapter 5; Stehman and Czaplewski 1998).

Alternatively, the cross-validation technique employed in the MRLC 2000 pilot study can be used to produce instant accuracy estimates. Whether these estimates are biased or not depends on the sampling design and spatial coverage of the training data points. Results from the MRLC 2000 pilot study suggest that such estimates might be unbiased if the training data set is collected following probability based sampling design and covers the entire study area. While inflated accuracies may result when significant spatial auto-correlations exist among the training samples (Campbell 1981), the cross-validation technique can, at the very least, provide users with preliminary information regarding the accuracy of the products they want to use.

The validity of classification accuracy estimates can be affected by two additional factors. One is possible labelling error of reference data points (Congalton and Green 1993), which may arise from errors of field crew members or photo interpreters, or from possible temporal discrepancies between reference data and satellite image. The other is location error that may exist between the reference data points and satellite image. Unfortunately, it is often difficult to quantitatively assess such errors and their impact on the validity of derived accuracy estimates.

Beyond classification

Landsat data can be used not only to develop forest land cover classifications, but also to estimate a suite of forest attributes, including tree canopy density, age, height, basal area, and tree bole diameter at breast height, among others (e.g. Cohen et al. 1995; Cohen et al. 2001). Information on such attributes at intermediate spatial resolutions is required for fire fuel modelling and many other forest management applications. In addition to the classifications developed in this Chapter, we have also estimated sub-pixel tree canopy density at the 30m resolution for the entire MRLC pilot study area using a regression tree technique. Shown to be robust for approximating complex non-linear relationships (Huang and Townshend 2002), the regression tree method was also used to estimate tree canopy density from ETM+ imagery in three two-scene areas in Virginia, Utah/Idaho, and Oregon (Huang et al. 2001). In these studies, the mean absolute difference and correlation coefficient (*r*) between predicted and actual tree canopy density were about 9 to 12 percent tree cover and 0.8 to 0.9, respectively.

6. CONCLUSIONS

Through two case studies, we have presented two approaches to extracting forest information from Landsat imagery in multi-scene regions. In both case studies, classifications were performed on multi-scene mosaics, which was more efficient and helped to achieve a higher degree of class consistency across multiple scenes than classifying single scenes individually. The two image pre-processing procedures – histogram bias adjustment and at-satellite reflectance method, were found effective for normalising the among-scene variations of clear and near cloud-free images in a semi-arid environment. Both the supervised decision tree classifier and the unsupervised approach produced satisfactory classification results. With adequate, well-distributed training data points readily available, the decision tree method should be more efficient and consistent than an unsupervised approach, especially for an area that needs to be remapped periodically. With cross-validation, the decision tree program can also generate instant accuracy estimates, which may be reasonably unbiased if the training points follow a probability based sampling design. This can be highly valuable, because statistically valid accuracy assessment over large areas is often very expensive. When only limited or no reference data points are available, however, the unsupervised approach may be more appropriate for extracting forest information at regional scales.

ACKNOWLEDGEMENTS

This work is made possible in part by the Raytheon Corporation under U.S. Geological Survey (USGS) contract 1434-CR-97-CN-40274. The FIA plot data were made available through an inter-agency agreement between USGS EROS Data Center (EDC) and the Forest Inventory and Analysis

(FIA) program of U.S. Forest Service's Rocky Mountain Research Station (RMRS). The authors are especially grateful to Ronald Tymcio and Larry DeBlander of RMRS for processing the FIA field plot data and photo-interpreted data for use in this study. The Fire Science Lab of RMRS and the Utah State University Remote Sensing Lab provided independent test data sets. The MRLC 2000 pilot study was supported by the MRLC 2000 mapping strategy team at EDC.

REFERENCES

Aplin, P., Atkinson, P. M., & Curran, P. J. (1997). Fine spatial resolution satellite sensors for the next decade. *International Journal of Remote Sensing*, 18, 3873-3881.

Arvidson, T., Gasch, J., & Goward, S. N. (2001). Landsat 7's long-term acquisition plan - An innovative approach to building a global imagery archive. *Remote Sensing of Environment*, 78, 13-26.

Bauer, E., & Kohavi, R. (1998). An empirical comparison of voting classification algorithms: bagging, boosting, and variants. *Machine Learning*, 5, 1-38.

Bauer, M. E., Burk, T. E., Ek, A. R., Coppin, P. R., Lime, S. D., Walsh, T. A., Walters, D. K., Befort, W., & Heinzen, D. F. (1994). Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering & Remote Sensing*, 60, 287-298.

Campbell, J. B. (1981). Spatial correlation effects upon accuracy of supervised classification of land cover. *Photogrammetric Engineering & Remote Sensing*, 47, 355-363.

Chan, J. C.-W., Huang, C., & DeFries, R. S. (2001). Enhanced algorithm performance for land cover classification using bagging and boosting. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 693-695.

Cihlar, J., Ly, H., & Xiao, Q. (1996). Land cover classification with AVHRR multichannel composites in northern environments. *Remote Sensing of Environment*, 58, 36-51.

Cohen, W. B., Fiorella, M., Bray, J., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering & Remote Sensing*, 64, 293-300.

Cohen, W. B., Maiersperger, T. K., Spies, T. A., & Oetter, D. R. (2001). Modelling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. *International Journal of Remote Sensing*, 22, 2279-2310.

Cohen, W. B., Spies, T. A., & Fiorella, M. (1995). Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A. *International Journal of Remote Sensing*, 16, 721-746.

Congalton, R., & Green, K. (1993). A practical look at sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing*, 59, 641-644.

Coppin, P. R., & Bauer, M. E. (1994). Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing*, 32, 918-927.

DeFries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. (1998). Global land cover classifications at 8km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168.

DeFries, R. S., & Townshend, J. R. G. (1993). Global land cover assessments: comparisons of ground-based data sets to classifications with AVHRR data. Goody, G., & Curran, P. (Eds.). *Environmental remote sensing from regional to global scales*. John Wiley and Sons, Chichester.

Edwards, T. C., Moisen, G. G., & Cutler, D. R. (1998). Assessing map accuracy in a remotely-sensed, ecoregion-scale cover-map. *Remote Sensing of Environment*, 63, 73-83.

Edwards, T. C., Scott, J. M., Homer, C. G., & Ramsey, R. D. (1993). Gap analysis: A geographic approach for assessing national biological diversity. *Natural Resources and Environmental Issues*, 2, 65-72.

Fleming, M. D., & Hoffer, R. M. (1979). Machine processing of Landsat MSS data and DMA topographic data for forest cover-type mapping. *Proceedings of the 1979 Machine Processing of Remotely Sensed Data Symposium*, 377-408. Purdue University/LARS, West Lafayette, Indiana.

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61, 399-409.

Friedl, M. A., Brodley, C. E., & Strahler, A. H. (1999). Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 969-977.

Fuller, R. M., Groom, G. B., & Jones, A. R. (1994a). The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering & Remote Sensing*, 60, 553-562.

Fuller, R. M., Groom, G. B., & Wallis, S. M. (1994b). The availability of Landsat TM images of Great Britain. *International Journal of Remote Sensing*, 15, 1357-1362.

Hall, F. G., Townshend, J. R., & Engman, E. T. (1995). Status of remote sensing algorithms for estimation of land surface state parameters. *Remote Sensing of Environment*, 51, 138-156.

Hansen, M., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331-1364.

Hansen, M., Dubayah, R., & DeFries, R. (1996). Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, 17, 1075-1081.

Henery, R. J. (1994). Methods for comparison. Michie, D., Spiegelhalte , D. J., & Taylor, C. C. (Eds.). *Machine learning, neural and statistical classification*, 107-130. Ellis Horwood, New York.

Homer, C., Huang, C., Yang, L., & Wylie, B. (2002). Development of a circa 2000 land cover database for the United States. *2002 ASPRS-ACSM Annual Conference and FIG XXII Congress*, Washington, D.C.

Homer, C. G., Ramsey, R. D., Edwards, T. C., & Falconer, A. (1997). Landscape cover-type modeling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering & Remote Sensing*, 63, 59-67.

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002a). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23, 725-749.

Huang, C., & Townshend, J. R. G. (2002). A stepwise regression tree for non-linear approximation: applications to estimating subpixel land cover. *International Journal of Remote Sensing*, in press.

Huang, C., Wylie, B., Homer, C., Yang, L., & Zylstra, G. (2002b). Derivation of a Tasseled cap transformation based on Landsat 7 at-satellite reflectance. *International Journal of Remote Sensing*, 23, 1741-1748.

Huang, C., Yang, L., Wylie, B., & Homer, C. (2001). A strategy for estimating tree canopy density using Landsat 7 ETM+ and high resolution images over large areas. *Third International Conference on Geospatial Information in Agriculture and Forestry*, Denver, Colorado, CD-ROM.

Hutchinson, C. F. (1982). Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing*, 48, 123-130.

Irish, R. R. (2000). Landsat 7 science data user's handbook. *Report 430-15-01-003-0*, NASA, http://ltpwww.gsfc.nasa.gov/IAS/handbook/handbook_toc.html.

Jensen, J. R. (1986). *Introductory digital image processing: a remote sensing perspective.*, Prentice-Hall, Englewood Cliffs, New Jersey.

Jones, B. T. & Smith, F. G. (2001). Mapping global land cover and changes economically and accurately: The GEOCOVER LC Approach. *Third International Conference on Geospatial Information in Agriculture and Forestry*. Denver, Colorado, CD-ROM.

Kimes, D. S., Nelson, R. F., Salas, W. A., & Skole, D. L. (1999). Mapping secondary tropical forest and forest age from SPOT HRV data. *International Journal of Remote Sensing*, 20, 3625-3640.

Loveland, T. R., & Shaw, D. M. (1996). Multiresolution land characterisation: building collaborative partnerships. *Proceedings of the ASPRS/GAP Symposium -- Gap analysis: a landscape approach to biodiversity planning*, 83-89. Charlotte, North Carolina.

Lunetta, R. S., & Balogh, M. E. (1999). Application of multi-temporal Landsat 5 TM imagery for wetland identification. *Photogrammetric Engineering & Remote Sensing*, 65, 1303-1310.

Lunetta, R. S., & Sturdevant, J. A. (1993). The North American landscape characterisation Landsat pathfinder project. *Pecora 12 Symposium: Land information from space-based systems*, 363-369. Sioux Falls, South Dakota.

Markham, B. L., & Barker, J. L. (1986). Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures. *EOSAT Landsat Technical Notes*, 1, 3-8.

McRoberts, R. E., Wendt, D. G., Nelson, M. D., & Hansen, M. H. (2002). Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, 81, 36-44.

Omernik, J. M. (1987). Ecoregions of the conterminous United States. *Annals of the Association of American Geographers*, 77, 118-125.

Pax-Lenney, M., Woodcock, C. E., Macomber, S. A., Gopal, S., & Song, C. (2001). Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment*, 77, 241-250.

Quinlan, J. R. (1993). *C4.5 programs for machine learning*, Morgan Kaufmann Publishers, San Mateo, California.

Richards, J. A. (1993). *Remote sensing digital image analysis: an introduction*. Springer-Verlag, Berlin.

Schott, J. R., Salvaggio, C., & Volghok, W. J. (1988). Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment*, 26, 1-16.

Scott, J. M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards, T. C., Ulliman, J., & Wright, R. G. (1993). Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs*, 123, 1-41.

Skole, D. L., Justice, C. O., Townshend, J. R. G., & Janetos, A. C. (1997). A land cover change monitoring program: strategy for an international effort. *Mitigation and Adaptation Strategies for Global Change*, 2, 157-175.

Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64, 331-344.

Sugumaran, R. (2001). Forest land cover classification using statistical artificial neural network approaches applied to IRS LISS - III sensor. *Geocarto International*, 16, 37-42.

Teillet, P. M., & Fedosejevs, G. (1995). On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing*, 21, 374.

Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles.* Addison-Wesley Publishing Company, London.

Townshend, J. R. G. (1992). Land cover. *International Journal of Remote Sensing*, 13, 1319-1328.

Tucker, C. J., Catlin, J. A., & Schneider, S. R. (1984). Monitoring vegetation in the Nile Delta with NOAA-6 and NOAA-7 AVHRR imagery. *Photogrammetric Engineering and Remote Sensing*, 50, 53-61.

USDA (1991). State Soil Geographic Data Base (STATSGO): Data users guide. *Report Miscellaneous Publication Number 1492.* U.S. Department of Agriculture, Soil Conservation Service, U.S. Government Printing Office.

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Driel, N. V. (2001). Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering & Remote Sensing*, 67, 650-662.

Vogelmann, J. E., Sohl, T., & Howard, S. M. (1998). Regional characterisation of land cover using multiple sources of data. *Photogrammetric Engineering & Remote Sensing*, 64, 45-57.

Yang, L., Homer, C., Hegge, K., Huang, C., & Wylie, B. (2001). A Landsat 7 Scene Selection Strategy for a National Land Cover Database. *IEEE International Geoscience and Remote Sensing Symposium.* Sydney, Australia, CD ROM, 1 disk.

Zhu, Z., & Evans, D. L. (1994). US forest types and predicted percent forest cover from AVHRR data. *Photogrammetric Engineering & Remote Sensing*, 60, 525-531.